Uncertainty in Prior Elicitations: a Nonparametric approach

Jeremy Oakley and Anthony O'Hagan

October 8, 2003

Global sensitivity analysis is recognized as an essential tool for investigating the effects of input parameter uncertainty in a complex model. To obtain meaningful results from a sensitivity analysis, it is important that the probability distributions for all the uncertain input parameters in the model accurately represent the beliefs of the model user or decision-maker. When little or no data related to these parameters are available, parameter distributions must be specified largely on the basis of expert knowledge. This is rarely a simple task.

A particular difficulty in this scenario is that to perform the global sensitivity analysis the full joint probability distribution is required for all the uncertain input parameters in the model. However, a full probability distribution implies an infinite number of probability judgments by the expert about the parameters, clearly something the expert is unable to provide. In practice it is only going to be possible to elicit a finite and typically small number of probability statements from the expert. These statements will typically take the form quantiles of the distribution, perhaps the mode and sometimes the mean or other moments. Such statements are not enough to identify the expert's probability distribution uniquely, and the usual approach is to fit some member of a convenient parametric family. There are two clear deficiencies in this solution. First, the expert's beliefs are forced to fit the parametric family. The parametric family may imply additional beliefs about the parameters that the expert does not agree with. Second, no account is then taken of the many other possible distributions that might have fitted the elicited statements equally well. This clearly has consequences for a global sensitivity analysis; other distributions might produce very different results when the uncertainty they are describe is propagated through the computer model under investigation.

We present an approach which tackles both of these deficiencies. Our model is nonparametric, allowing the expert's distribution to take any continuous form. It also quantifies the uncertainty in the resulting elicited distribution. Formally, the expert's density function is treated as an unknown function, about which we make inference. The result is a posterior distribution for the expert's density function. The posterior mean serves as a 'best fit' elicited distribution, while the variance around this fit expresses the uncertainty in the elicitation.

Specifically, this is achieved by using a Gaussian process to describe our own beliefs about the expert's distribution. Our prior specification contains proper prior beliefs about the smoothness of the expert's distribution, but is ultimately vague in that we do not include any of our own beliefs about likely values of the uncertain input parameter. Data then comes in the form of the expert's summaries, such as their mean and various quantiles. Properties of Gaussian processes can then be exploited to update our beliefs about the expert's distribution analytically, conditional on various hyperparameters in our Gaussian process model. Finally, Markov Chain Monte Carlo methods are used to remove the conditioning on these hyperparameters to give a full, probabilistic description of our uncertainty about the expert's distribution.

Illustrations of our method are given using some simple real elicitation exercises.